

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

JC551 U.S.PTO
09/386824
08/31/99

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出願年月日
Date of Application:

1999年 6月28日

出願番号
Application Number:

平成11年特許願第181959号

願人
Applicant(s):

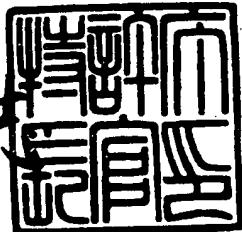
富士通株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

1999年 7月 8日

特許庁長官
Commissioner,
Patent Office

伴佐山 建志



【書類名】 特許願
【整理番号】 9900622
【提出日】 平成11年 6月28日
【あて先】 特許庁長官殿
【国際特許分類】 H03M 7/00
【発明の名称】 音声符号化方法、音声符号化装置、及び音声復号装置
【請求項の数】 10
【発明者】
【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通
株式会社内
【氏名】 大田 恒士
【発明者】
【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通
株式会社内
【氏名】 鈴木 政直
【発明者】
【住所又は居所】 福岡県福岡市博多区博多駅前三丁目22番8号 富士通
九州ディジタル・テクノロジ株式会社内
【氏名】 土永 義照
【特許出願人】
【識別番号】 000005223
【氏名又は名称】 富士通株式会社
【代理人】
【識別番号】 100074099
【住所又は居所】 東京都千代田区二番町8番地20 二番町ビル3F
【弁理士】
【氏名又は名称】 大菅 義之
【電話番号】 03-3238-0031

【選任した代理人】

【識別番号】 100067987

【住所又は居所】 神奈川県横浜市港北区太尾町1418-305（大倉
山二番館）

【弁理士】

【氏名又は名称】 久木元 彰

【電話番号】 045-545-9280

【先の出願に基づく優先権主張】

【出願番号】 平成10年特許願第246724号

【出願日】 平成10年 9月 1日

【手数料の表示】

【予納台帳番号】 012542

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9705047

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声符号化方法、音声符号化装置、及び音声復号装置

【特許請求の範囲】

【請求項1】 音源符号語を複数の非零振幅値のみで構成した符号帳を用いる、合成による分析型のベクトル量子化に基づく音声符号化方法であって、前記各非零振幅値のサンプル位置を、インデックスと、音声の特徴量を表す伝送パラメータを用いて可変制御する過程を含むことを特徴とする音声符号化方法。

【請求項2】 音源符号語を複数の非零振幅値のみで構成した符号帳を用いる、合成による分析型のベクトル量子化に基づく音声符号化方法によって符号化された音声信号を復号する音声復号方法であって、

前記各非零振幅値のサンプル位置を、インデックスと、音声の特徴量を表す伝送パラメータを用いて可変制御する過程を含むことを特徴とする音声復号方法。

【請求項3】 請求項1又は2の何れか1項に記載の方法であって、前記各非零振幅値のサンプル位置を、前記インデックスと、前記音声の特徴量を表す伝送パラメータであるピッチ周期に対応するラグ値を用いて可変制御する過程を含むことを特徴とする音声符号化方法又は音声復号方法。

【請求項4】 請求項1又は2の何れか1項に記載の方法であって、前記各非零振幅値のサンプル位置を、前記インデックスと、前記音声の特徴量を表す伝送パラメータであるピッチ周期に対応するラグ値及びピッチゲイン値とを用いて可変制御する過程を含むことを特徴とする音声符号化方法又は音声復号方法。

【請求項5】 請求項3又は4の何れか1項に記載の方法であって、前記各非零振幅値のサンプル位置を、前記ラグ値と前記音声の符号化単位であるフレーム長との関係に応じて前記ラグ値に対応する区間内で再構成する過程を含むことを特徴とする音声符号化方法又は音声復号方法。

【請求項6】 請求項4に記載の方法であって、前記非零振幅値のサンプル位置を、前記ピッチゲイン値に応じて前記ラグ値に

対応する区間内で再構成する過程を含むことを特徴とする音声符号化方法又は音声復号方法。

【請求項7】 音源符号語を複数の非零振幅値のみで構成した符号帳を用いる、合成による分析型のベクトル量子化に基づく音声符号化装置であって、前記各非零振幅値のサンプル位置を、インデックスと、音声の特徴量を表す伝送パラメータを用いて可変制御する構成可変符号帳手段を含む、ことを特徴とする音声符号化装置。

【請求項8】 音源符号語を複数の非零振幅値のみで構成した符号帳を用いる、合成による分析型のベクトル量子化に基づく音声符号化装置によって符号化された音声信号を復号する音声復号装置であって、前記各非零振幅値のサンプル位置を、インデックスと、音声の特徴量を表す伝送パラメータを用いて可変制御する構成可変符号帳手段を含む、ことを特徴とする音声復号装置。

【請求項9】 請求項7又は8の何れか1項に記載の装置であって、前記構成可変符号帳手段は、前記各非零振幅値のサンプル位置を、前記インデックスと、前記音声の特徴量を表す伝送パラメータであるピッチ周期に対応するラグ値を用いて可変制御する、ことを特徴とする音声符号化装置又は音声復号装置。

【請求項10】 請求項7又は8の何れか1項に記載の装置であって、前記構成可変符号帳手段は、前記各非零振幅値のサンプル位置を、前記インデックスと、前記音声の特徴量を表す伝送パラメータであるピッチ周期に対応するラグ値及びピッチゲイン値とを用いて可変制御する、ことを特徴とする音声符号化装置又は音声復号装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、A-b-S (Analysis-by-Synthesis; 合成による分析) 型ベクトル量子化を用いる音声符号化／復号技術に関する。

【0002】

【従来の技術】

CELP (Code Excited Liner Prediction; 符号励振線形予測) 符号化方式に代表される A-b-S 型ベクトル量子化による音声符号化方式は、PCM 音声信号の伝送レートを例えば 64 k bits/sec (キロビット/秒) から 4~16 k bits/sec 程度に圧縮する場合に適用されている。このような音声符号化方式は、企業内通信システムやディジタル移動無線システム等において音声品質を保持しながら情報圧縮を行うものとして、その要望が高いものである。

【0003】

図13は、従来のA-b-S型ベクトル量子化方式の説明図である。51は符号帳、52は係数器、53は線形予測合成フィルタ、54は減算器、55は誤差電力評価部を示す。

【0004】

A-b-S型ベクトル量子化符号器では、まず、係数器52が、符号帳51から読み出されたコードベクトルCにゲインgを乗算する。次に、線形予測合成フィルタ53が、上記乗算結果を入力し、再生信号gACを出力する。続いて、減算器54が、入力信号Xから上記再生信号gACを減算することにより、それらの差分である誤差信号Eを出力する。更に、誤差電力評価部55が、誤差信号Eに基づいて誤差電力を算出する。以上の処理が、符号帳51内の全てコードベクトルCと複数種類のゲインgに対して実行され、上記誤差電力が最小となるコードベクトルCのインデックスとゲインgとが算出され、それらが復号器に伝送される。

【0005】

A-b-S型ベクトル量子化復号器では、まず、符号器から伝送されてきたインデックスに対応するコードベクトルCが、符号帳51から読み出される。次に、係数器52が、上記コードベクトルCに、符号器から伝送されてきたゲインgを乗算する。そして、線形予測合成フィルタ53が、上記乗算結果を入力し、復号された再生信号gACを出力する。復号器では、減算器54と誤差電力評価部55は不要である。

【0006】

以上のようにして、A-b-S型ベクトル量子化の符号器においては、コードベクトルCに対して合成（復号）処理が実行されながら分析が行われることが特徴である。

【0007】

図14は、上述のA-b-S型ベクトル量子化方式に基づく、従来の代表的なCELP方式の説明図である。

このCELP方式では、符号帳として、周期（ピッチ）性音源に対応する適応符号帳と、雑音（ランダム）性音源に対応する固定符号帳の2種類が備えられる。そして、それぞれの符号帳に基づいて、主に周期性音声（有聲音等）のためのA-b-S型ベクトル量子化処理と、それに続く主に雑音性音声（無聲音又は背景音等）のためのA-b-S型ベクトル量子化処理とが、順次実行されることが特徴である。

【0008】

図14において、61は固定符号帳、62は適応符号帳、63及び64は係数器、65及び66は線形予測合成フィルタ、67及び68は誤差電力評価部、69及び70は減算器である。ランダム性音源に対応する固定符号帳61と、ピッチ性音源に対応する適応符号帳62は、それぞれメモリによって構成されている。また、係数器63及び64と線形予測合成フィルタ65及び66と誤差電力評価部67及び68と減算器69及び70は、DSP（デジタル・シグナル・プロセッサ）等の演算素子によって実現することができる。

【0009】

上記構成を有するCELP符号器において、適応符号帳62、係数器64、線形予測合成フィルタ66、減算器70、及び誤差電力評価部68からなる部分は、周期性音声に対して有効な伝送パラメータを出力する。また、Pは適応符号帳から出力される適応コードベクトル、bは係数器64におけるゲイン、Aは線形予測合成フィルタ66の伝達特性である。

【0010】

この部分における符号化処理は、図13で説明した、符号帳51、係数器52

、線形予測合成フィルタ53、減算器54、及び誤差電力評価部55による符号化処理と原理は同様である。但し、適応符号帳62内のサンプルは、過去の励起信号が帰還されることにより、適応的に変化する。復号器についても、図13で説明した、符号帳51、係数器52、及び線形予測合成フィルタ53による復号処理と同様の処理が実行される。但しこの場合も、適応符号帳62内のサンプルは、過去の励起信号が帰還されることにより、適応的に変化する。

【0011】

一方、固定符号帳61、係数器63、線形予測合成フィルタ65、減算器69、及び誤差電力評価部67からなる部分は、減算器70が入力信号Xから線形予測合成フィルタ66より出力される最適な再生信号bAPを減算することにより出力する雑音性信号X'に対して有効な伝送パラメータを出力する。この部分における符号化処理も、図13で説明した、符号帳51、係数器52、線形予測合成フィルタ53、減算器54、及び誤差電力評価部55による符号化処理と原理は同じである。この場合、固定符号帳61には、予め固定サンプルが格納される。復号器についても、図13で説明した、符号帳51、係数器52、及び線形予測合成フィルタ53による復号処理と同様の処理が実行される。

【0012】

ここで、固定符号帳61は、予め固定サンプル値に対応する雑音コードベクトルCを格納するものである。従って、例えば、ベクトル次元長を40（サンプリング周波数が8kHzである場合における5ms（ミリ秒）の期間内のサンプル数に相当する）、基本ベクトル数を1024とすると、固定符号帳61は、40k（キロ）ワードのメモリ容量を必要とすることになる。

【0013】

即ち、固定符号帳61が、全てのサンプル値を独立に格納するためには、膨大なメモリ容量が必要となり、CELP音声コーデックの実現上の大きな問題となっていた。

【0014】

そこで、この問題点を解決する目的で、少数の非零サンプル値が固定位置に配置されることにより代数的解法で音源探索処理が可能とされるACELP（A1

gebraic Code Excited Linear Prediction) 方式が提案されている (J. P. Adoul 他 "Fast CELP coding based on algebraic codes" Proc. IEEE International conference on acoustics speech and signal processing, pp. 1957-1960 (April, 1987) 参照)。

【0015】

図15は、代数符号帳を用いた従来のACELP方式の構成図である。代数符号帳71は図14の固定符号帳61に対応し、係数器72は図14の係数器63に対応し、線形予測合成フィルタ73は図14の線形予測合成フィルタ65に対応し、減算器74は図14の減算器69に対応し、誤差電力評価部75は図14の誤差電力評価部67に対応する。図15におけるA-b-S処理では、図13又は図14で説明した処理と同様に、インデックス*i*に対応して代数符号帳71から生成されるコードベクトル C_i とゲイン g とを用いて、A-b-S処理が実行される。

【0016】

このACELP方式では、非零サンプルの振幅値と位置に制約が加えられることにより、演算量及びメモリ量の大幅な削減が実現されている。このとき、例えば図16に示されるように、コードベクトル C_0, C_1, \dots, C_{M-1} を格納したN次元、Mサイズの代数符号帳71が用意されるが、フレーム内の非零サンプルの個数が固定化されかつ非零サンプルが等間隔に配置されるという制約から、各コードベクトル C_0, C_1, \dots, C_{M-1} を代数的解法により生成することが可能となる。図16の例では、4個の非零サンプル i_0, i_1, i_2, i_3 の各サンプル位置が規格化され、その振幅値が±1.0とされる。この4サンプル位置以外の他のサンプル位置の振幅は、零とされる。

【0017】

また図16に示す代数符号帳71の右側に示されるように、 i_0, i_1, i_2, i_3 に対応するコードベクトルのサンプル値パターンは、例えば、コードベクトル C_0 に対応するパターン(0, ..., 0, +1, 0, ..., 0, -1, 0, ...)

$\dots, 0, +1, 0, \dots, 0, -1, 0, \dots$ のように、振幅が零であるサンプル位置を除く、振幅が±1であるサンプル位置が i_0, i_1, i_2, i_3 に従って決定される。即ち4個の非零サンプルと $N - 4$ 個の零サンプルとの合計 N サンプルを要素としたコードベクトルについて、4個の非零サンプル i_n ($n = 0, 1, 2, 3$) のそれぞれは、振幅情報（振幅の絶対値が1に固定されて極性のみを示す）のための1ビットと、 2^k 通りの候補の中の1つを指定する位置情報 m_n のための K ビットとの、合計 $K + 1$ ビットで表すことができる。

【0018】

非零サンプルの位置は、ITU-T（国際電気通信連合の電機通信標準化部門）のG. 729又はG. 723. 1によって標準化されている。

例えば標準規格G. 729に対応した図16の表77において、1フレームに対応する40サンプルについて、非零サンプル $i_0 \sim i_2$ の各位置情報 $m_0 \sim m_2$ は、それぞれ8か所の候補を有するため、それぞれその中の1か所を指定するために3ビットで表現することができる。また、非零サンプル i_3 の位置情報 m_3 は、16か所の候補を有するため、その中の1か所を指定するために4ビットで表現することができる。また、非零サンプル $i_0 \sim i_3$ の各振幅情報 $s_0 \sim s_3$ は、各振幅の絶対値が1.0で固定され、その極性が表現されればよいから、それぞれ1ビットで表現することができる。以上より、G. 729では非零サンプル $i_0 \sim i_3$ は、図16の76として示されるように、それぞれ1ビットで構成される振幅情報 $s_0 \sim s_3$ と、それぞれ3ビット又は4ビットで構成される位置情報 $m_0 \sim m_3$ とからなる、17ビットのデータにより構成することができる。

【0019】

また、標準規格G. 723. 1に対応した図16の表78においては、非零サンプル $i_0 \sim i_3$ の各位置候補が、各非零サンプル間で1サンプルおきになるよう决定される。これにより、非零サンプル $i_0 \sim i_3$ の各位置情報 $m_0 \sim m_3$ は、それぞれ3ビットで表現することができる。非零サンプル $i_0 \sim i_3$ の各振幅情報 $s_0 \sim s_3$ は、標準規格G. 729の場合と同様に、それぞれ1ビットで表現することができる。以上より、G. 723. 1では非零サンプル $i_0 \sim i_3$

は、図16の76として示されるように、それぞれ1ビットで構成される振幅情報 $s_0 \sim s_3$ と、それぞれ3ビットで構成される位置情報 $m_0 \sim m_3$ とからなる、16ビットのデータにより構成することができる。

【0020】

ここで例えば、第*i*番目の符号語が $s^i_{n_0}, m^i_{n_0}$ (但し、 $n=0, 1, 2, 3$) の値を待つ時に、符号語サンプル $c^i(n)$ は、次式により定義することができる。

【0021】

【数1】

$$c^i(n) = s^i_0 \delta(n - m^i_1) + s^i_1 \delta(n - m^i_1) \\ + s^i_2 \delta(n - m^i_2) + s^i_3 \delta(n - m^i_3)$$

ここで、 s^i_n は非零サンプルの振幅情報、 m^i_n は非零サンプルの位置情報である。また、 $\delta()$ はデルタ関数であり、

$$\delta(n) = 1 \quad \text{for } n = 0$$

$$\delta(n) = 0 \quad \text{for } n \neq 0$$

である。

【0022】

また、誤差電力 E^2 は、図15に示される入力信号 X と、ゲイン g と、コードベクトル C_i と、線形予測合成フィルタ73のインパルス応答の行列 H とを用いて、次式で表現できる。

【0023】

【数2】

$$E^2 = (X - g H C_i)^2$$

この誤差電力 E^2 を最小とするための評価関数 $\arg\max(F_i)$ は、次式で表現できる。

【0024】

【数3】

$$\arg\max(F_i) = [(X^T H C_i)^2 / ((H C_i)^T (H C_i))]$$

ここで、

【0025】

【数4】

$$X^T H = D = d \quad (i)$$

【0026】

【数5】

$$H^T H = \Phi = \phi \quad (i, j)$$

とすると、数3式に示される評価関数 $\text{argmax}(F_i)$ は、次式で表現することができる。

【0027】

【数6】

$$\text{argmax}(F_i) = [(D^T C_i)^2 / \{ (C_i)^T \Phi C_i \}]$$

なお、大文字はベクトルを示す。

【0028】

前述の数4式及び数5式には、コードベクトル C_i の要素が含まれないから、符号語のパターン数（サイズ）Mが多い場合でも、予め計算しておくことができる。従って、数6式は、数3式に比較して、高速な演算が可能となる。

【0029】

また、コードベクトル C_i に関する処理は、前述のように、振幅が ±1.0 の4サンプルについて行うもので、数6式の分母と分子とは、それぞれ以下の数7式及び数8式で示される演算となる。

【0030】

【数7】

$$(D^T C_i)^2 = \{ \sum_{i=0}^3 s_i d(m_i) \}^2$$

【0031】

【数8】

$$(C_i)^T \Phi C_i = \sum_{i=0}^3 \phi(m_i, m_i) + 2 \sum_{i=0}^2 \sum_{j=i+1}^3 s_i s_j \phi(m_i, m_j)$$

なお、 $\sum_{i=0}^3$ は、 $i = 0$ から $i = 3$ までの累算を示す。

【0032】

上記数7式及び数8式の演算量は、パラメータ（次元数）Nには依存せずかつ演算量も少ないため、符号語パターン数Mに対応する回数だけ演算が実行されたとしても演算量は膨大にはならない。従って、図15に示される代数符号帳71を用いる構成では、図14に示される固定符号帳61を用いる構成に比較して、大幅に演算量を低減することができる。また、代数符号帳71から出力される各コードベクトルは、振幅情報（極性情報）と位置情報とから代数的に生成することができるから、メモリに各コードベクトルを記憶させる必要がなく、メモリ量の大幅な削減が可能となる。

【0033】

【発明が解決しようとする課題】

前述のACELP方式では、メモリ量の削減及び演算量の削減を図ることができるものであるが、フレーム内の非零サンプルの本数が4本に固定され、かつサンプル位置が等間隔になるような制約が加えられているため、フレーム長と非零サンプル数との2つのパラメータにより符号語インデックスを表現するビットレートが決定され、符号語インデックスを表現するために比較的多くのビットが必要となるという問題点を有している。

【0034】

例えばITU-Tの標準規格G.729において、1フレームを40サンプルとしたとき、図16の表77に示すように、符号語インデックスとして合計17ビットが使用されることになる。このビット数は、G.729が規定する総伝送容量(8 k bits/sec, 80 bits/10 msec)の42%に相当することになる。

【0035】

また1フレームを80サンプルとすると、非零サンプルの位置情報を表現するために必要なビット数は、前述の場合よりそれぞれ1ビット分増加する。このため、符号語インデックスとして合計21ビットが使用されることになる。このビット数は、G.729が規定する総伝送容量の62.5%にも相当することになり、1フレーム40サンプルの場合より大幅に増加する。

【0036】

一般に4 k b i t s / s e c 程度の極低ビットレート音声C O D E Cを実現するためには、フレーム長の拡張が必須となる。しかし、このような要求に上述した従来のA C E L P方式を適用すると、符号語インデックスの伝送ビットレートの大幅な増加が問題となってくるのである。即ち、従来のA C E L P方式は、伝送効率を上げて単位時間あたりのパラメータ伝送ビット数を削減することにより低ビットレート化を図るという要求を、阻害してしまうという問題点を有している。

【0037】

この問題と共に、従来のA C E L P方式は、フレーム長が拡張されたときに、フレーム長以下のピッチ周期に対する同定能力が低下するという問題も有している。

【0038】

本発明の課題は、非零振幅値のみで構成される音源符号語を用いるA-b-S型ベクトル量子化に基づく音声符号化／復号方式において、符号語インデックスの伝送量の一定化とピッチ周期に対する同定能力の維持を図ることにある。

【0039】

【課題を解決するための手段】

本発明は、（1）音源符号語を複数の非零振幅値のみで構成した符号帳を用いる、合成による分析型のベクトル量子化に基づく音声符号化技術であって、非零振幅値のサンプル位置を、インデックスと、音声の特徴量を表す伝送パラメータを用いて可変制御するものである。この場合の伝送パラメータとして、ピッチ周期に対応するラグ値を用いることができる。更に、ピッチゲイン値を用いることができる。また、ラグ値の大小関係又はピッチゲイン値に対応して、非零振幅値のサンプル位置をラグ値に対応する区間内で再構成するように構成することができる。

【0040】

【発明の実施の形態】

以下、図面を参照しながら本発明の実施の形態について詳細に説明する。

図1及び図2は本発明の原理説明図であり、1及び1'は構成可変符号帳、2

及び $2'$ は係数器、 3 及び $3'$ は線形予測合成フィルタ、 4 は減算器、 5 は誤差電力評価部を示す。

【0041】

構成可変符号帳 1 及び $1'$ は、例えば複数の非零サンプル値からなるコードベクトルを出力する代数符号帳に相当するものであるが、インデックス i と、ピッチ周期（ラグ値）等の伝送パラメータ p とに基づいて、非零サンプルの位置を制御することにより、自身を再構成する機能を有する。このとき、構成可変符号帳 1 及び $1'$ は、非零サンプル数を変化させることなく、非零サンプル位置を可変制御する。これにより、符号語インデックスの伝送に必要なビット数の増加を抑制することが可能となる。

【0042】

図1に示される本発明の原理構成の符号器では、まず、係数器 2 が、インデックス i と伝送パラメータ p とに基づいて非零サンプルの位置が制御された上で構成可変符号帳 1 から出力されるコードベクトル C_i に、ゲイン g を乗算する。次に、線形予測合成フィルタ 3 が、上記乗算結果を入力し、再生信号 gAC_i を出力する。続いて、減算器 4 が、入力信号 X から上記再生信号 gAC_i を減算することにより、それらの差分である誤差信号 E を出力する。更に、誤差電力評価部 5 が、誤差信号 E に基づいて誤差電力を算出する。以上の処理が、構成可変符号帳 1 から出力される全てコードベクトル C_i と複数種類のゲイン g に対して実行され、上記誤差電力が最小となるコードベクトル C_i のインデックス i とゲイン g とが算出され、それらが復号器に伝送される。

【0043】

図2に示される本発明の原理構成の復号器ではまず、パラメータ分離部 6 が、符号器から伝送されてきた受信データから各パラメータを分離する。次に、構成可変符号帳 $1'$ は、上記分離されたパラメータのうちのインデックス i と伝送パラメータ p に基づいて、コードベクトル C_i を出力する。次に、係数器 $2'$ が、上記コードベクトル C_i に、パラメータ分離部 6 で分離されたゲイン g を乗算する。そして、線形予測合成フィルタ $3'$ が、上記乗算結果を入力し、復号された再生信号 gAC を出力する。なお、特には図示しないが、線形予測合成フィルタ

3'には、パラメータ分離部6から、線形予測パラメータが与えられる。

【0044】

図1及び図2の構成における伝送パラメータpとしては、音声信号の特性に対応して種々選択することが可能であり、例えば、ピッチ周期（ラグ値）やゲイン等を採用することができる。

【0045】

図3及び図4は上述した図1及び図2の原理構成に基づく本発明の第1の実施の形態の説明図であり、11及び11'は構成可変符号帳、12及び12'は係数器、13及び13'は線形予測合成フィルタ、14は減算器、15は誤差電力評価部、16は非零サンプル位置制御部、17はピッチ強調フィルタ、18はパラメータ分離部を示す。

【0046】

構成可変符号帳11及び11'は、図3中下部（図4も同様）に示されるように、インデックスiと伝送パラメータであるピッチ周期（ラグ値）l（Lの小文字）とを入力する非零サンプル位置制御部16と、この非零サンプル位置制御部16の出力信号とピッチ周期（ラグ値）l（エル）とを入力するピッチ強調フィルタ17とから構成される。非零サンプル位置制御部16は、非零サンプル数は変化させないが、ピッチ周期（ラグ値）l（エル）によって非零サンプルの位置を可変制御する。ピッチ強調フィルタ17は、ラグ値がフレーム長よりも短い場合に、ラグ値に対応する長さ以上のサンプルを過去のラグ値から合成するための帰還フィルタである。

【0047】

なお、図3及び図4に示される各部の機能は、DSP（デジタル・シグナル・プロセッサ）等の演算素子によって実現することもできる。

従来のACELP方式では、フレーム長に応じてその範囲全体に格納されるように非零サンプルが割り当てられていた。しかし、ピッチ周期に対応するラグ値がフレーム長よりも短い場合に、ラグ値に対応する長さ以上のサンプルは帰還フィルタを使って過去のラグ値から合成するような構成を採用することができる。このような場合に、非零サンプルを、フレーム内のラグ値に対応する範囲より広

い範囲にわたって割り当てるのは、無駄である。

【0048】

そこで、本実施の形態では、非零サンプル位置制御部16が、まず非零サンプルをラグ値の範囲内でのみ割り当てる。これと同時に、同制御部16は、ラグ値がフレーム長の半分に対応する値を越えるような場合には、ラグ値の範囲内に割り当てられる非零サンプルのうち、ピッチ強調フィルタ17による帰還処理の影響が少ないので後半部分に割り当てられる非零サンプルを間引いてその位置を可変制御する。これにより、ラグ値及びフレーム長が変化しても、非零サンプル数を一定数に保つことが可能となって、符号語インデックスの伝送に必要なビット数の増加を抑制することが可能となる。

【0049】

まず、図3及び図4の第1の実施の形態の構成の全体的な動作は、図1及び図2で説明した原理構成の動作と同様である。

図5は、図3及び図4の構成可変符号帳11及び11'内に構成される非零サンプル位置制御部16が実行する演算処理を示すフローチャートである。以下、1フレーム=80サンプル(8kHzサンプリング)、非零サンプル数=4、ラグ(lag)値=20サンプル(400Hz)~147サンプル(54.4Hz)、インデックス伝送ビット=17ビットとした場合を例として説明する。

【0050】

まず、非零サンプル位置の初期化が行われる(図5のステップA1)。ここでは、40要素からなる配列データ smp_pos[i] ($0 \leq i < 40$)に、等間隔の非零サンプル位置 $i = 0 \sim 39$ がセットされる。

【0051】

次に、入力するピッチ周期に対応するラグ値が判定される。ここで、ラグ値は、図3及び図4では特に図示しないが、ACELP処理の前段の処理である適応符号帳を用いたA-b-S処理(図14の上半分の構成に対応)によって算出されている。

【0052】

まず、ラグ値が第1の設定値:40以下か否かが判定され(図5のステップA

2)、この判定がYESならば、図5のステップA6が実行されることにより、各非零サンプル位置が登録される。

【0053】

この結果、ピッチ周期に対応するラグ値が40以下ならば、非零サンプル位置が、図6(a)に示されるように決定される。この配置は、前述したITU-T標準規格G.729に対応する図16の表77に示されるものと同一である。

【0054】

一方、図5のステップA2の判定がNOならば、ラグ値が第2の設定値：80以上か否かが判定され(図5のステップA3)、この判定がNOの場合には、図5のステップA5の非零サンプル位置の制御処理におけるforループ処理によって配列データsmp_pos[]の内容が順次変更され、その後その変更された配列データを使って、ステップA6の非零サンプル位置の登録処理が実行される。

【0055】

この結果、ピッチ周期に対応するラグ値が40より大きく80よりも小さい場合、例えば45である場合には、非零サンプル位置が例えば図6(b)に示されるように決定される。この配置は、図7の説明図に示されるように、図6(a)の表で示される配置において、サンプル位置35、37、39が省略されその代わりにサンプル位置40、42、44が追加されたものである。

【0056】

より具体的には、例えばラグ値=45の場合に、初期値として、 $i_x = 0$, $i_y = 0$ とされ、 $(lag - 41) / 2 + 1 = 3$ となるから、3個のサンプル位置について位置制御が行われる。即ち、 $smp_pos[39 - iy] = ix$ の演算により、まず、 $i_x = 40$, $i_y = 0$ であるから、サンプル位置データsmp_pos[39]にサンプル位置39の代わりにサンプル位置40が設定され、次に、 $i_x += 2$, $i_y += 2$ により $i_x = 42$, $i_y = 2$ となるから、サンプル位置データsmp_pos[37]にサンプル位置37の代わりにサンプル位置42が設定され、更に、 $i_x = 44$, $i_y = 4$ となるから、サンプル位置データsmp_pos[35]にサンプル位置35の代わりにサンプル位置

4.4 設定される。

【0057】

このように、本実施の形態では、ピッチ周期に対応するラグ値が40より大きく80よりも小さい場合には、ラグ値が40から増加した分に応じたサンプル数だけサンプル位置が間引かれてラグ値の範囲内で再構成されることにより、非零サンプル数を変更することなくその位置が再構成される。

【0058】

続いて、図5のステップA3の判定がYESであった場合には、図5のステップA4のクリッピング処理が実行される。即ち、ラグ値がフレーム長に対応する80を越えた場合には、フレーム長の範囲外にまで非零サンプルを割り当てることは無意味であるため、ラグ値が80にクリップされた上で、図5のステップA5の非零サンプルの位置制御処理とそれに続くステップA6の非零サンプル位置の登録処理が実行される。この結果、非零サンプル位置が例えば図6(c)に示されるように決定される。

【0059】

以上の制御処理により、ラグ値が増加した場合でも、ラグ値に対応して非零サンプル位置が再構成されることにより、非零サンプル数を変更しないで符号語インデックスの必要伝送ビット数を17ビットのままとすることが可能となる。

【0060】

図8は、図3及び図4の構成可変符号帳11及び11'を構成するピッチ強調フィルタ17が行うピッチ強調処理の説明図であり、31及び34は係数器、32は加算器、33は遅延回路を示す。

【0061】

図8において、係数器31及び34と加算器32と遅延回路33とを含む構成部分の伝達関数は、 $P(z) = \alpha / (1 - \beta z^{-1\text{ag}})$ となる。なお、 α は係数器31の係数、 β は係数器34の係数、 1ag はラグ値を示す。例えば、係数器31の係数 α は、 $0 \sim (1\text{ag} - 1)$ の範囲では、 $\alpha = 1.0$ 、 $1\text{ag} \sim 79$ の範囲では $\alpha = 0.0$ とし、係数器34の係数 β は 1.0 とすることができる。なお、係数 α 、 β はこれらの値に限定されるものではなく、他の値を設定することも

もちろん可能である。

【0062】

上記構成を有する回路構成により、ラグ値がフレーム長よりも短い場合に、フレーム内のラグ値に対応する長さ以上のサンプルが過去のラグ値から帰還させられて合成される。この結果、ピッチ周期に同期した系列生成を行うことができ、ピッチ同定能力の低下を回避することができる。

【0063】

図9及び図10は前述の図1及び図2の原理構成に基づく本発明の第2の実施の形態の説明図であり、21及び21'は構成可変符号帳、22及び22'は係数器、23及び23'は線形予測合成フィルタ、24は減算器、25は誤差電力評価部、26は非零サンプル位置制御部、27はピッチ同期フィルタ、28はパラメータ分離部を示す。

【0064】

上述の図9及び図10の第2の実施の形態の構成の全体的な動作は、図1及び図2で説明した原理構成の動作と同様である。

構成可変符号帳21及び21'は、本発明の第1の実施の形態に対応する構成符号帳11及び11'（図3及び図4）の場合と同様に、非零サンプル位置制御部26とピッチ同期フィルタ27とから構成される。この第2の実施の形態の構成が第1の実施の形態の構成と異なる点は、非零サンプル位置制御部26とピッチ同期フィルタ27が、伝送パラメータとして、ピッチ周期に対応するラグ値1（エル）に加えてピッチゲインGも入力している点である。

【0065】

適応符号帳を用いたA-b-S処理（図14の上半分の構成に対応）によって算出されるピッチ周期に対応したラグ値としては、入力音声が明確なピッチ周期を有しない場合でも、探索範囲内で最も可能性の高い値が選択される。従って、雑音性音源が適当であると言われている無聲音や背景音の区間でも、擬似的なピッチ周期が抽出され、このピッチ周期の情報が符号器から復号器へ伝送される。この場合に、ピッチゲインGが大きいことはピッチ周期性が大きいことを示しており、ピッチゲインGが小さいことは無聲音や背景音などのようにピッチ周期性

が小さいことを示している。そこで、本発明の第2の実施の形態では、伝送パラメータの1つとして、ピッチゲインGが採用されるものである。

【0066】

図11は、図9及び図10の構成可変符号帳21及び21'内に構成される非零サンプル位置制御部26が実行する演算処理を示すフローチャートである。このフローチャートにおいて、ステップB1、B3、B4、B7、B5、及びB6の各制御処理は、この順でそれぞれ、本発明の第1の実施の形態に対応する図5のフローチャートのステップA1、A2、A3、A4、A5、及びA6と同じ処理である。

【0067】

第2の実施の形態が第1の実施の形態と異なる点は、ピッチゲインGが閾値よりも小さい場合の処理である。即ち、図11のステップB2において、ピッチゲインGが閾値よりも小さいか否かが判定され、その判定がYESの場合には、ピッチ周期の設定が意味をなさないため、ステップB7においてラグ値がフレーム長と同じ80にクリップされた上で、第1の実施の形態の場合と同様の処理が実行される。

【0068】

以上の制御処理により、更なる特性向上を図ることが可能となる。

図12は、入力音声X（図中A、図14のXに対応）と、本実施の形態への雑音性入力信号X'（図中B、図1等のX'に対応）と、本発明における構成可変符号帳（図1の1等）からの出力信号の各波形例を示す図である。

【0069】

以上、本発明の各実施の形態について説明したが、本発明は、前述の各実施の形態のみに限定されるものではなく、種々付加変更することができる。例えば、フレーム長、サンプル数等は適用システムに対応して任意に選定することができる。また、伝送パラメータとしては、例えば音声のホルマントに対応するようなものを用いてもよい。更に、本発明は、ACELP方式のみでなく、非零サンプルを複数用い、その非零サンプル位置を伝送パラメータによって制御する音声符号化方式に適用できるものである。

【0070】

【発明の効果】

本発明によれば、A-b-S型ベクトル量子化において符号帳から出力される非零サンプルの位置を、インデックスと、ラグ値やピッチゲイン等の音声の特徴量を表す伝送パラメータを用いて、非零サンプル数を増加させることなく変更制御することができる。この結果、本発明は、フレーム長の拡張によっても所要伝送ビット数を増加する必要がなく、伝送効率の低下を回避できるという利点を有する。

【0071】

また、本発明は、ピッチ同定能力をピッチ強調処理等によって容易に確保できるという利点を有する。

【図面の簡単な説明】

【図1】

本発明の原理説明図（符号化探索処理）である。

【図2】

本発明の原理説明図（復号側再生処理）である。

【図3】

本発明の第1の実施の形態の説明図（符号化探索処理）である。

【図4】

本発明の第1の実施の形態の説明図（復号側再生処理）である。

【図5】

本発明の第1の実施の形態のフローチャートである。

【図6】

本発明の実施の形態のラグ値による構成可変符号帳の説明図である。

【図7】

本発明の実施の形態のラグ値に応じた非零サンプル位置の説明図である。

【図8】

ピッチ強調処理の説明図である。

【図9】

本発明の第2の実施の形態の説明図（符号化探索処理）である。

【図10】

本発明の第2の実施の形態の説明図（復号側再生処理）である。

【図11】

本発明の第2の実施の形態のフローチャートである。

【図12】

各信号の波形例を示す図である。

【図13】

従来のA-b-S型ベクトル量子化の説明図である。

【図14】

従来のCELP方式の説明図である。

【図15】

従来のACELP方式の構成図である。

【図16】

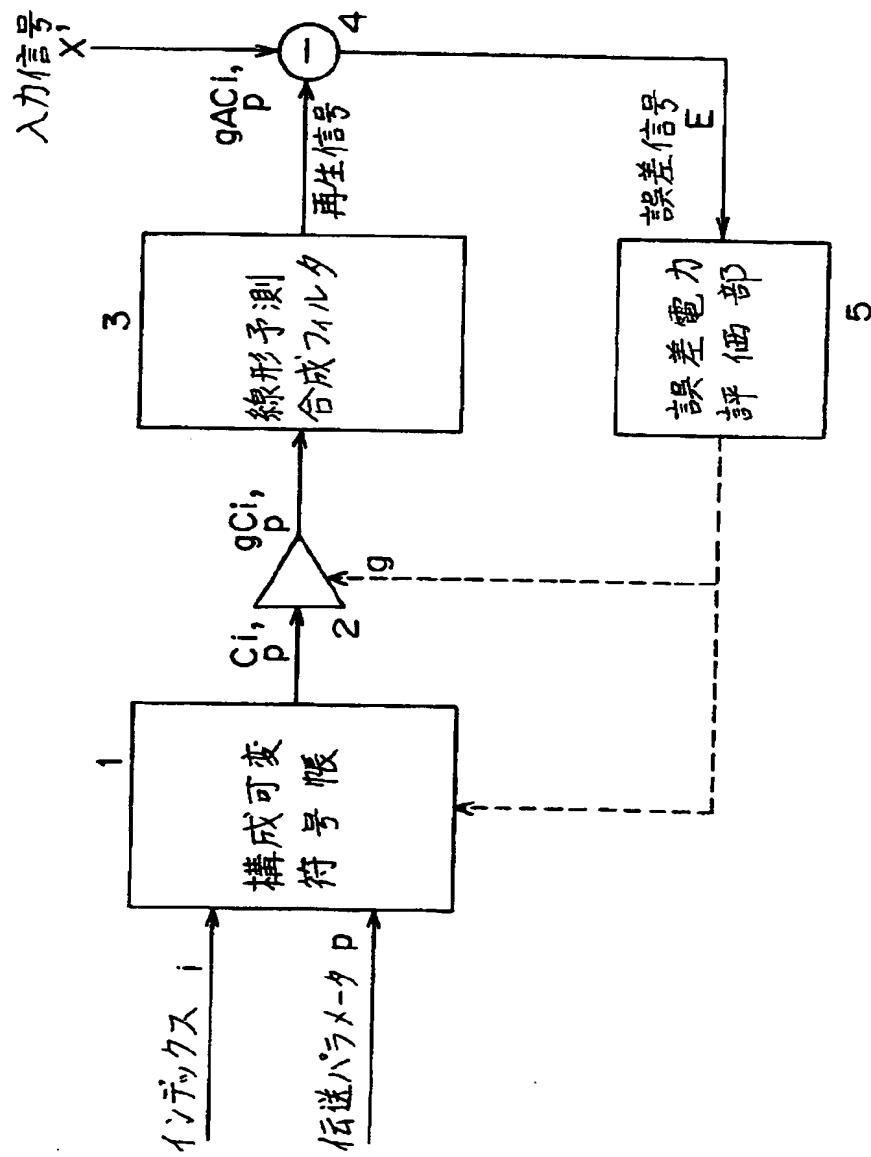
ACELP方式の概要の説明図である。

【符号の説明】

- 1、 1' 構成可変符号帳
- 2、 2' 係数器
- 3、 3' 線形予測合成フィルタ
- 4 減算器
- 5 誤差電力評価部

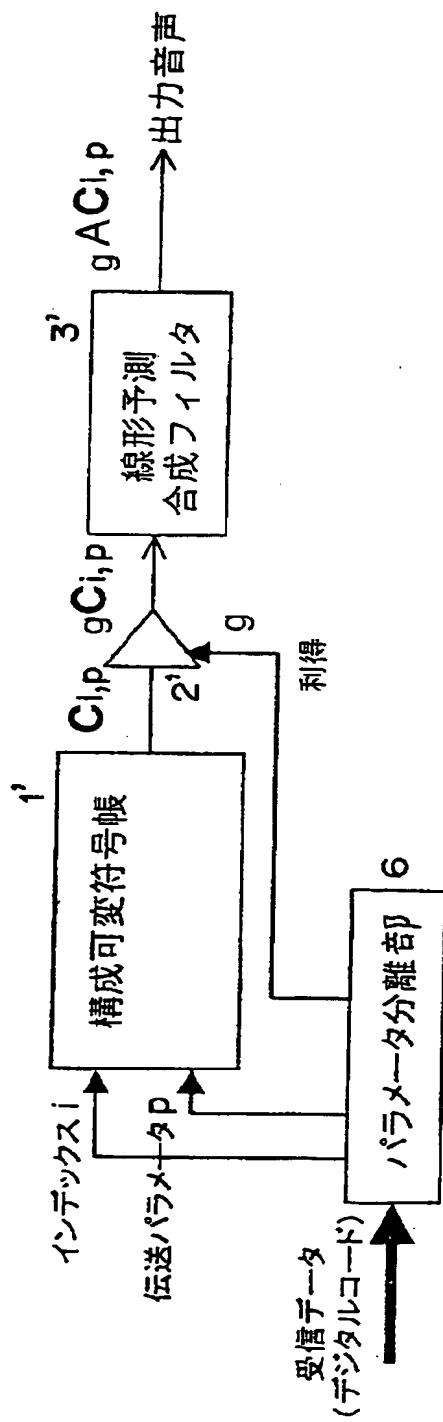
【書類名】 図面
【図1】

本発明の原理説明図(符号化探索処理)



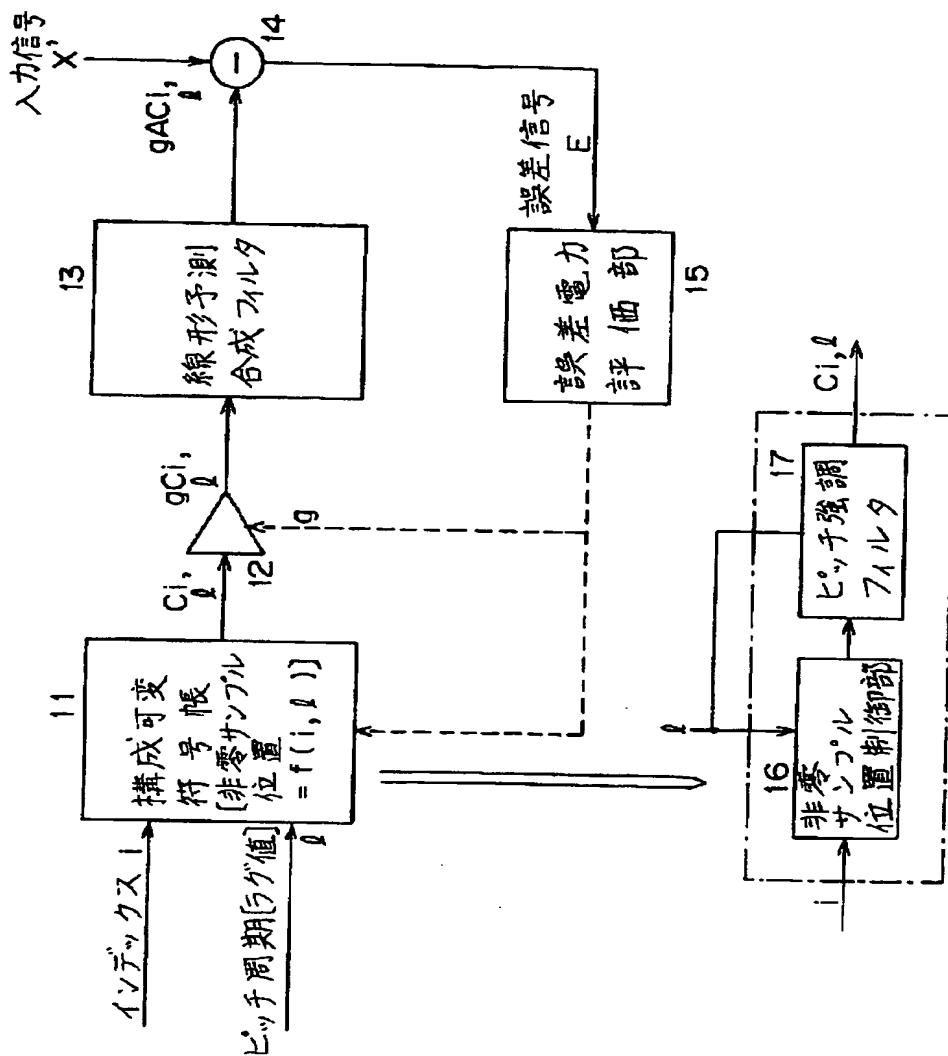
【図2】

本発明の原理説明図(復号側再生処理)



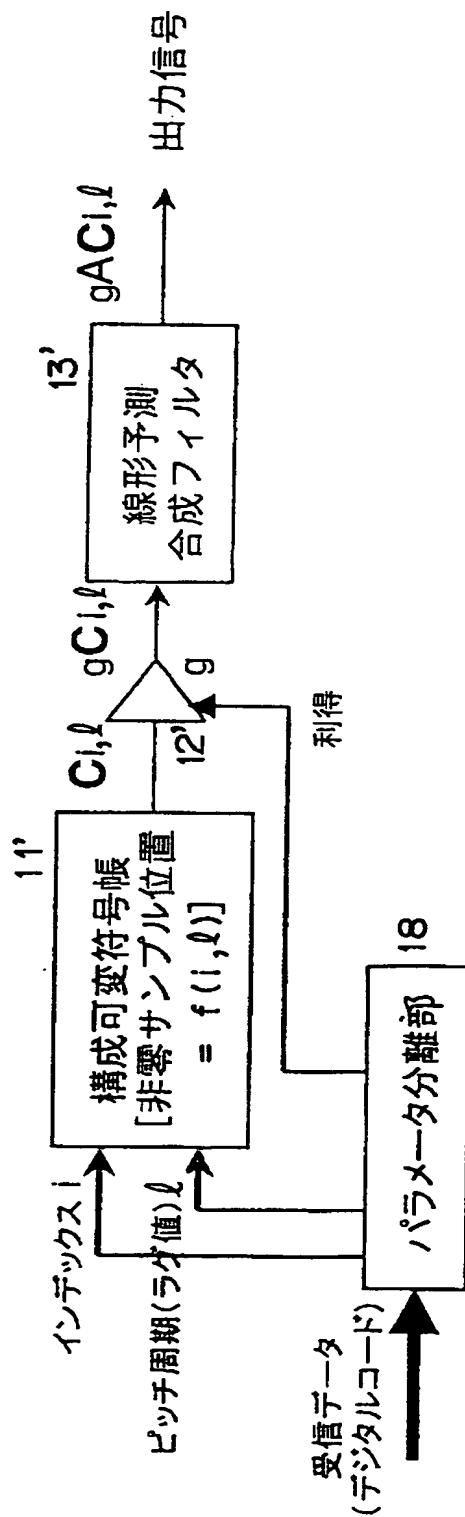
【図3】

本発明の第1の実施の形態の説明図
(符号化探索処理)



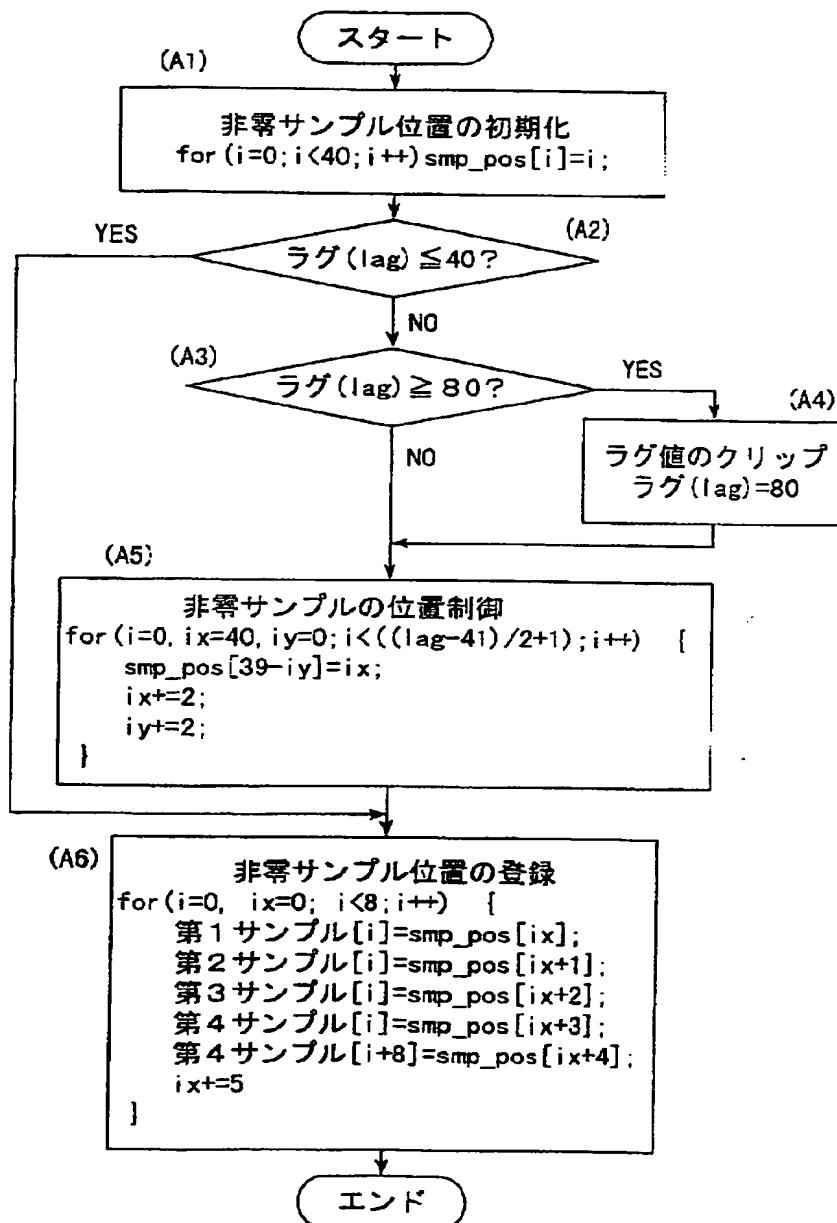
【図4】

本発明の第1の実施の形態の説明図
(復号側再生処理)



【図5】

本発明の第1の実施の形態のフローチャート



【図6】

本発明の実施の形態のラグ値による構成可変符号帳の説明図

20≤ラグ値≤40

(a)

	非零サンプル位置								
i ₀	0	5	10	15	20	25	30	35	
i ₁	1	6	11	16	21	26	31	36	
i ₂	2	7	12	17	22	27	32	37	
i ₃	3	8	13	18	23	28	33	38	
	4	9	14	19	24	29	34	39	

40<ラグ値<80(ラグ値=45の場合)

(b)

	非零サンプル位置								
i ₀	0	5	10	15	20	25	30	44	
i ₁	1	6	11	16	21	26	31	36	
i ₂	2	7	12	17	22	27	32	42	
i ₃	3	8	13	18	23	28	33	38	
	4	9	14	19	24	29	34	40	

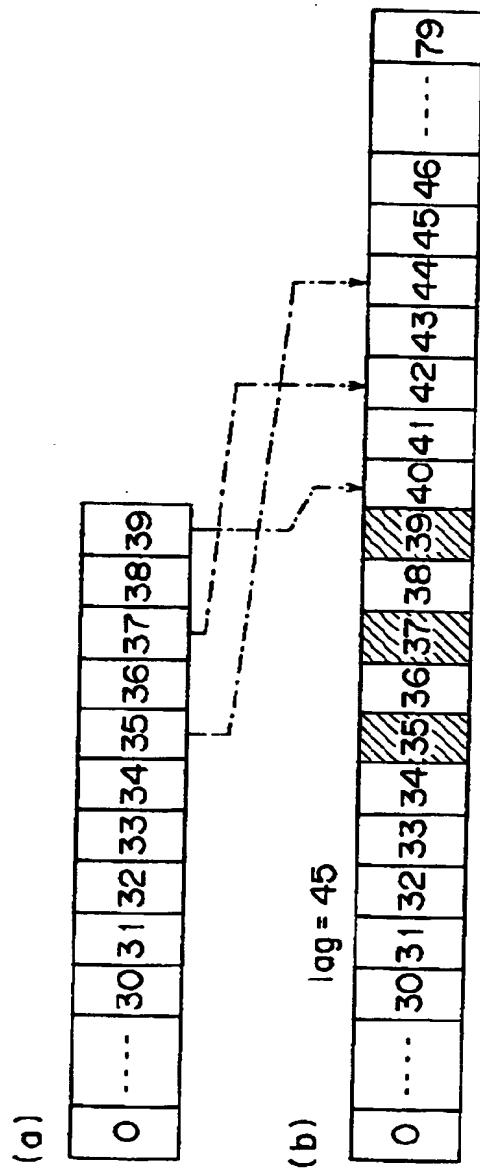
80≤ラグ値

(c)

	非零サンプル位置								
i ₀	0	74	10	64	20	54	30	44	
i ₁	78	6	68	16	58	26	48	36	
i ₂	2	72	12	62	22	52	32	42	
i ₃	76	8	66	18	56	28	46	38	
	4	70	14	60	24	50	34	40	

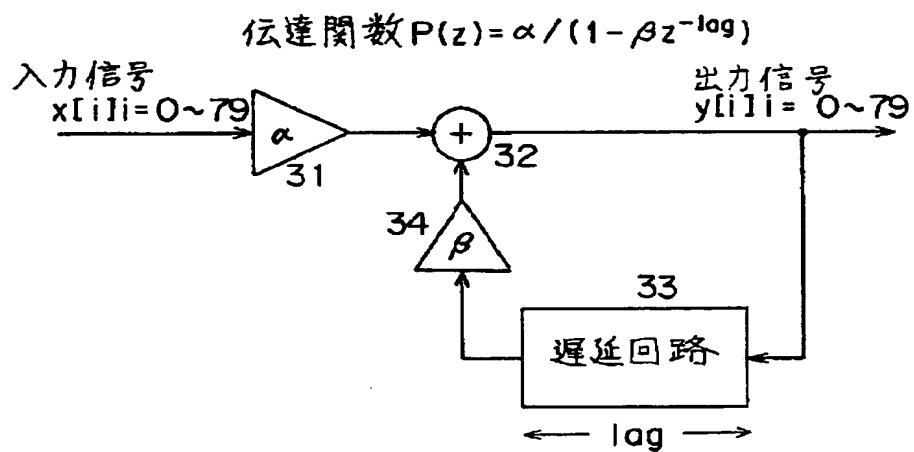
【図7】

本発明の実施の形態のラグ値に応じた
非零サンプル位置の説明図



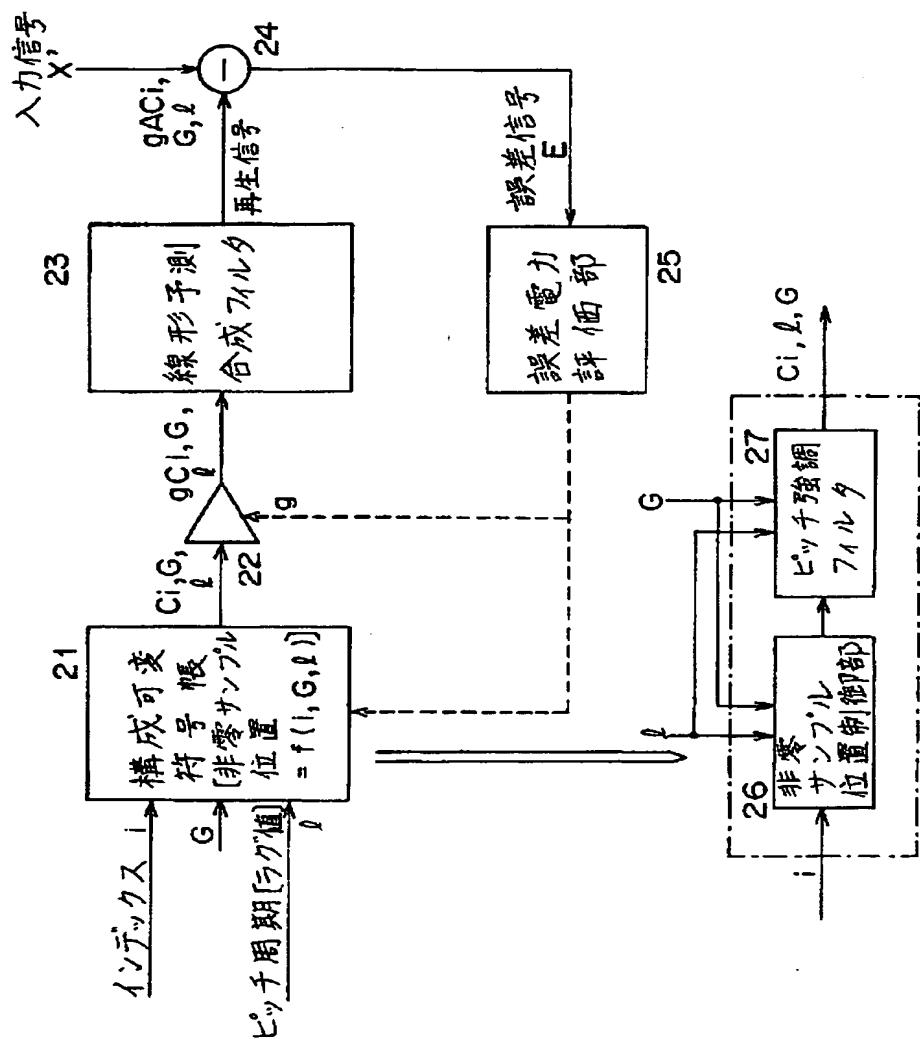
【図8】

ビーチ強調処理の説明図



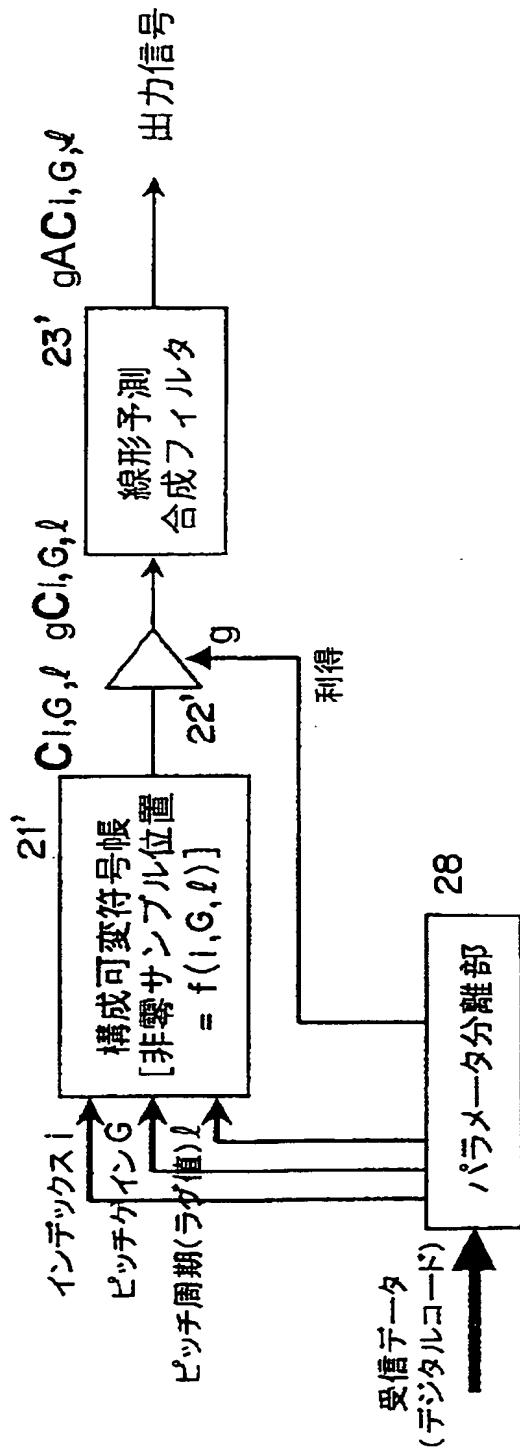
【図9】

本発明の第2の実施の形態の説明図
(符号化探索処理)



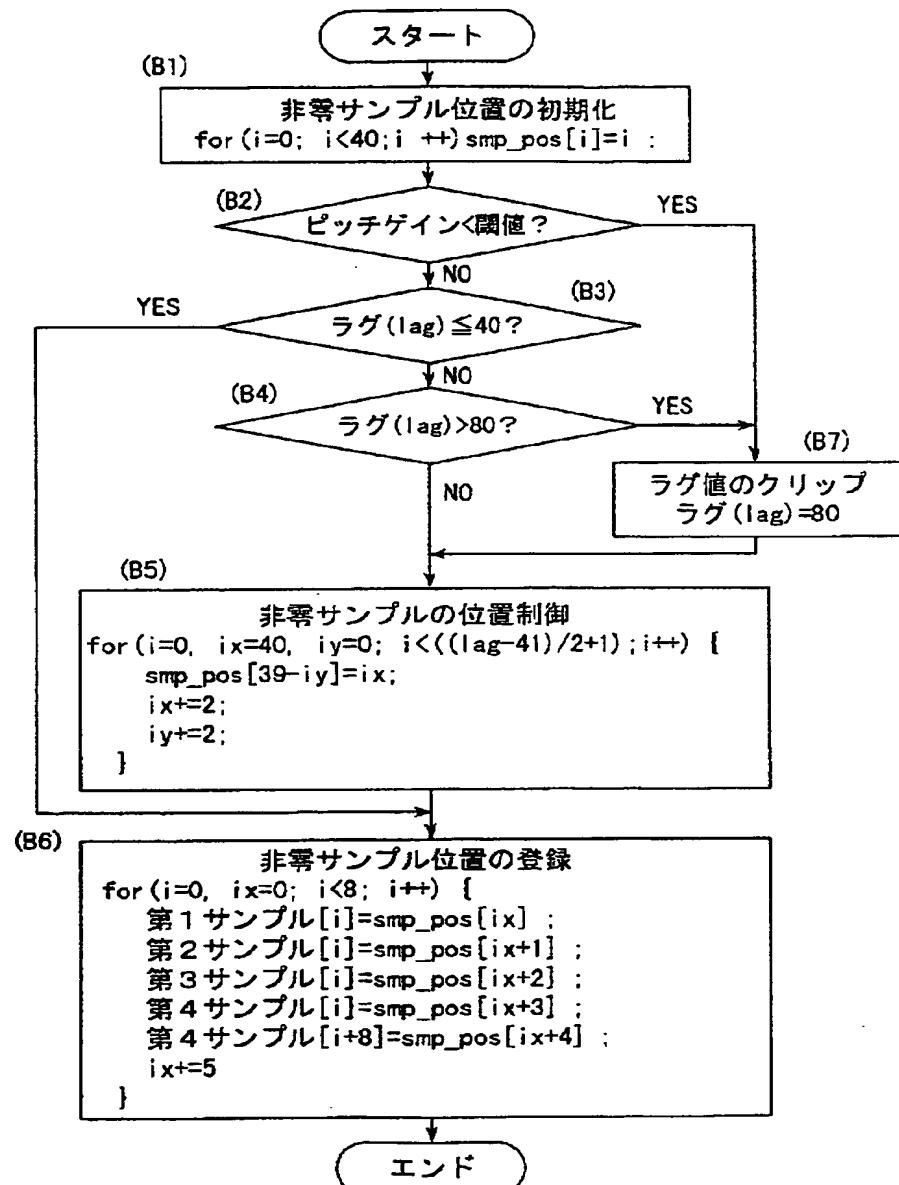
【図10】

本発明の第2の実施の形態の説明図
(復号側再生処理)



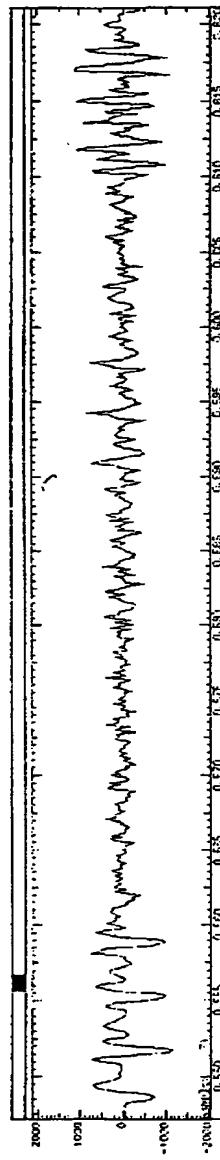
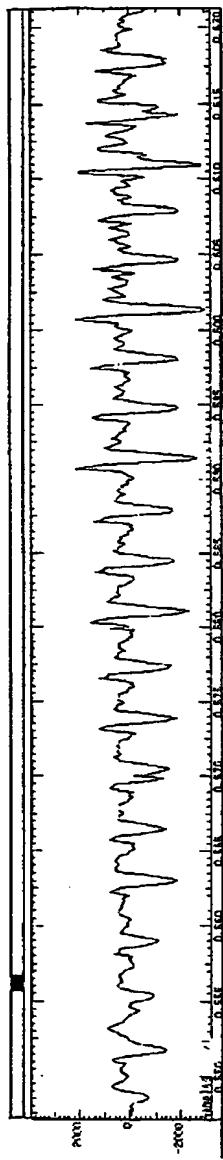
【図11】

本発明の第2の実施の形態のフローチャート



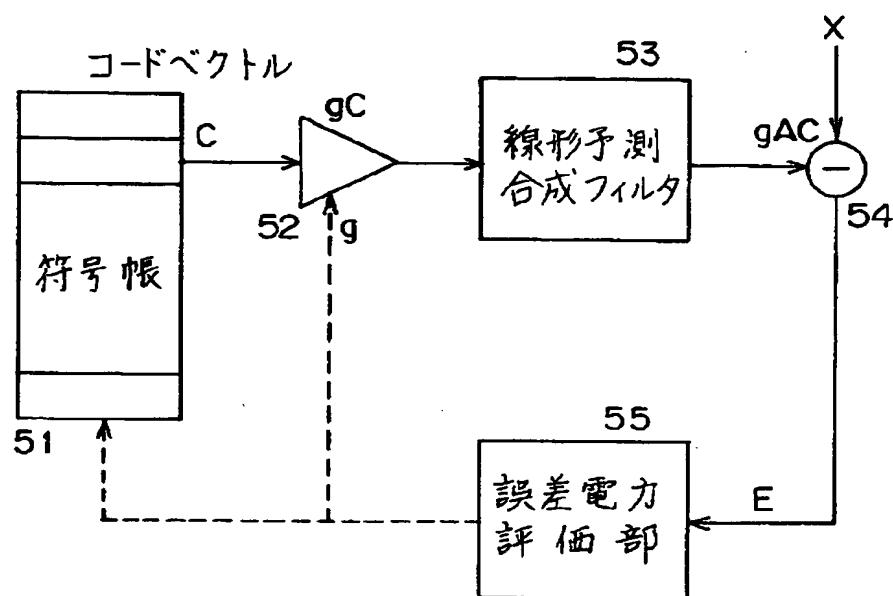
【図12】

各信号の波形例を示す図



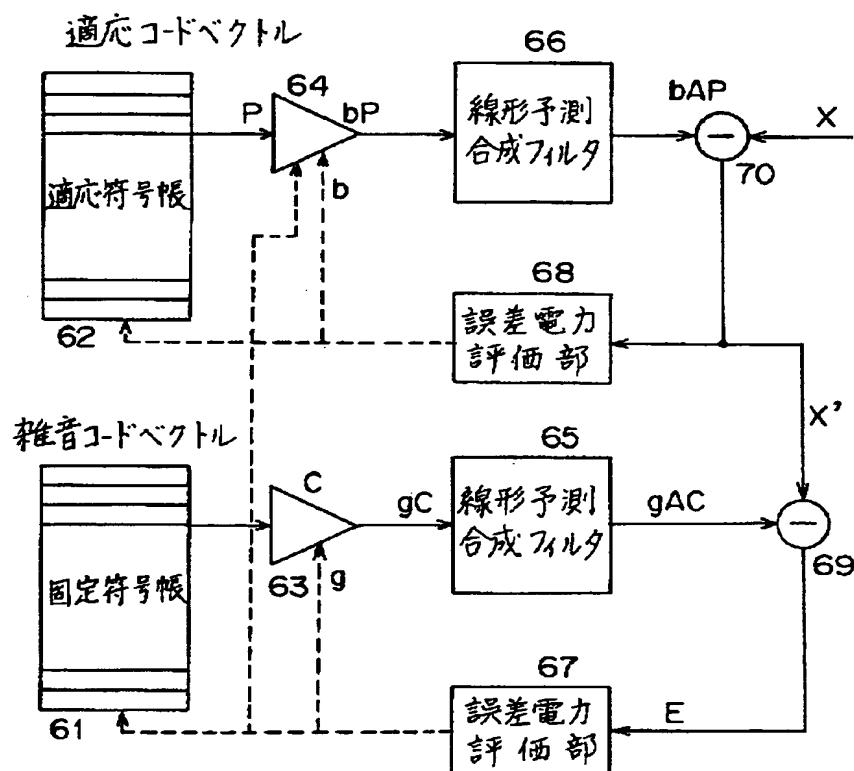
【図13】

従来のA-b-S型ベクトル量子化の説明図



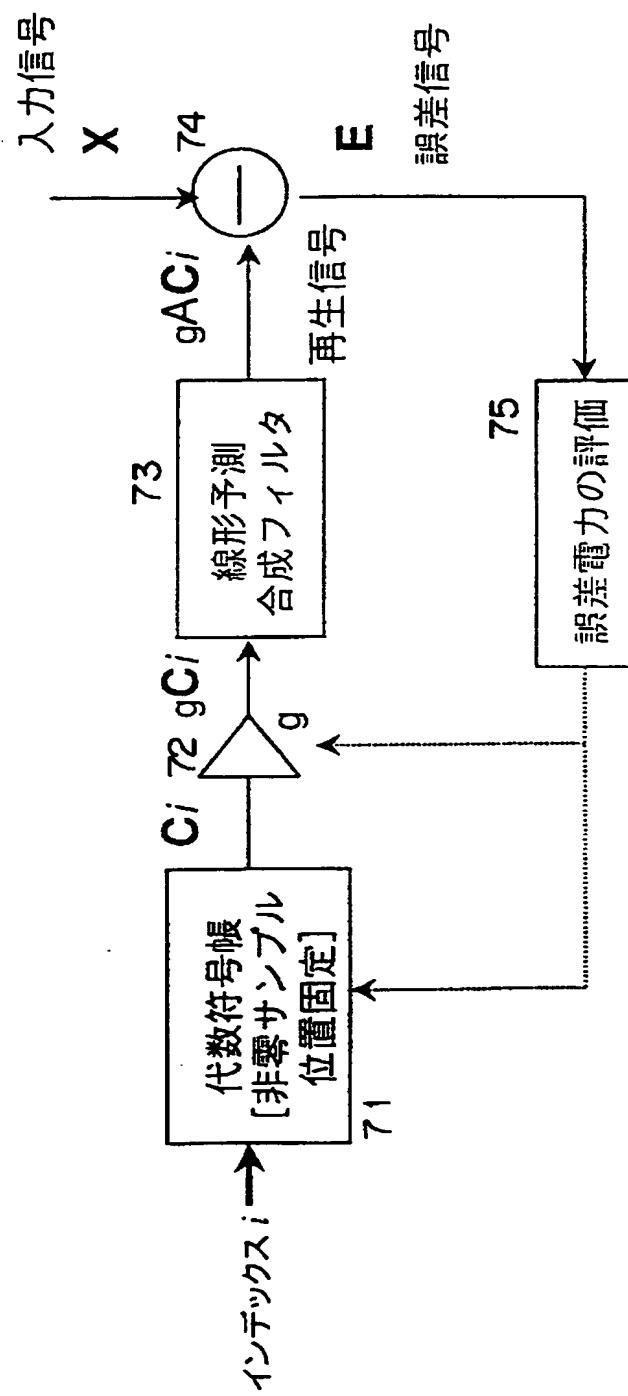
【図14】

従来のCELP方式の説明図



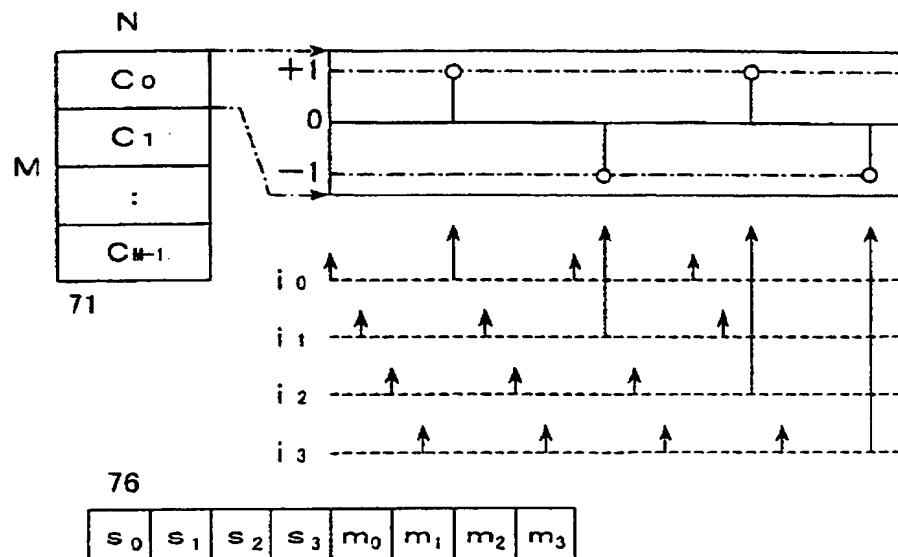
【図15】

従来の ACELP 方式の構成図



【図16】

A C E L P 方式の概要の説明図



77 符号 位置 ピット数

i_0	s_0	m_0	0, 5, 10, 15, 20, 25, 30, 35	3
i_1	s_1	m_1	1, 6, 11, 16, 21, 26, 31, 36	3
i_2	s_2	m_2	2, 7, 12, 17, 22, 27, 32, 37	3
i_3	s_3	m_3	3, 8, 13, 18, 23, 28, 33, 38 4, 9, 14, 19, 24, 29, 34, 39	4

合計17ビット

78 符号 位置 ピット数

i_0	s_0	m_0	0, 8, 16, 24, 32, 40, 48, 56	3
i_1	s_1	m_1	2, 10, 18, 26, 34, 42, 50, 58	3
i_2	s_2	m_2	4, 12, 20, 28, 36, 44, 52, (60)	3
i_3	s_3	m_3	6, 14, 22, 30, 38, 46, 54, (62)	3

合計16ビット

【書類名】 要約書

【要約】

【課題】 非零振幅値のみで構成される音源符号語を用いるA-b-S型ベクトル量子化に基づく音声符号化／復号方式において、符号語インデックスの伝送量の一定化とピッチ周期に対する同定能力の維持を図ることを目的とする。

【解決手段】 係数器2が、インデックス*i*と伝送パラメータ*p*とに基づいて非零サンプルの位置が制御された上で構成可変符号帳1から出力されるコードベクトル*C_i*に、ゲイン*g*を乗算する。線形予測合成フィルタ3は、上記乗算結果を入力し、再生信号*g A C_i*を出力する。減算器4は、入力信号Xから上記再生信号*g A C_i*を減算することにより、誤差信号Eを出力する。誤差電力評価部5は、誤差信号Eに基づいて誤差電力を算出する。以上の処理が、全てコードベクトル*C_i*とゲイン*g*に対して実行され、上記誤差電力が最小となるコードベクトル*C_i*のインデックス*i*とゲイン*g*とが算出され、復号器に伝送される。

【選択図】 図1

出願人履歴情報

識別番号 [000005223]

1. 変更年月日 1996年 3月26日

[変更理由] 住所変更

住 所 神奈川県川崎市中原区上小田中4丁目1番1号

氏 名 富士通株式会社